Lesson 2: Deterministic finite state automata

Theme: Deterministic finite state automata.

A deterministic finite state automaton (DFA) is a system $\mathcal{A} = \langle \Sigma, Q, q_0, F, \delta \rangle$, where each component is as follows.

- $\Sigma$ is the alphabet.
- $Q$ is a finite set of states.
- $q_0 \in Q$ is the initial state.
- $F \subseteq Q$ is the set of final states.
- $\delta : Q \times \Sigma \rightarrow Q$ is the transition function.

Remark 2.1 A DFA $\mathcal{A} = \langle \Sigma, Q, q_0, F, \delta \rangle$ can be visualised as a directed graph as follows.

- The vertices are elements of $Q$.
- There is an edge from $p$ to $p'$ labeled with $a$, if $\delta(p, a) = p'$.

On input word $w = a_1 \cdots a_n$, the run of $\mathcal{A}$ on $w$ is the sequence:

$$p_0 \ a_1 \ p_1 \ a_2 \ p_2 \ \cdots \ a_n \ p_n,$$

where $p_0 = q_0$ and $\delta(p_i, a_{i+1}) = p_{i+1}$, for each $i = 0, \ldots, n - 1$.

Sometimes we are interested in a run that does not start from the initial state. In that case, we can define the run of $\mathcal{A}$ on $w$ starting from state $q$ as the sequence defined as above, but with condition $p_0 = q$. That is,

$$p_0 \ a_1 \ p_1 \ a_2 \ p_2 \ \cdots \ a_n \ p_n,$$

where $p_0 = q$ and $\delta(p_i, a_{i+1}) = p_{i+1}$, for each $i = 0, \ldots, n - 1$.

A run is called an accepting run, if $p_0 = q_0$ and $q_n \in F$. We say that $\mathcal{A}$ accepts $w$, if there is an accepting run of $\mathcal{A}$ on $w$. The language of all words accepted by $\mathcal{A}$ is denoted by $L(\mathcal{A})$.

A language $L$ is called a regular language, if there is a DFA $\mathcal{A}$ such that $L(\mathcal{A}) = L$.

Remark 2.2 Let $\mathcal{A} = \langle \Sigma, Q, q_0, F, \delta \rangle$ be a DFA.

- The empty string $\varepsilon$ is accepted by $\mathcal{A}$ if and only if $q_0 \notin F$.
- For every word $w$, there is exactly one run of $\mathcal{A}$ on $w$.

Theorem 2.3 Regular languages are closed under boolean operations, i.e., intersection, union, and complementation. More formally, it can be stated as follows.

- For every DFA $\mathcal{A}$, there is a DFA $\mathcal{A}'$ such that $L(\mathcal{A}') = \Sigma^* - L(\mathcal{A})$.
- For every two DFA $\mathcal{A}_1$ and $\mathcal{A}_2$, there is a DFA $\mathcal{A}'$ such that $L(\mathcal{A}') = L(\mathcal{A}_1) \cap L(\mathcal{A}_2)$.
- For every two DFA $\mathcal{A}_1$ and $\mathcal{A}_2$, there is a DFA $\mathcal{A}'$ such that $L(\mathcal{A}') = L(\mathcal{A}_1) \cup L(\mathcal{A}_2)$.